



June 10, 2025

**VIA ECF**

Hon. Ona T. Wang  
Daniel Patrick Moynihan United States Courthouse  
500 Pearl St.  
New York, NY 10007-1312

**Re: In re: OpenAI, Inc., Copyright Infringement Litigation, 1:25-md-03143-SHS-OTW; this document relates to The Center for Investigative Reporting, Inc. v. OpenAI, Inc. et al., Case No. 1:24-cv-04872-SHS-OTW**

Dear Judge Wang:

OpenAI respectfully requests a pre-motion conference to address The Center for Investigative Reporting (“CIR”)’s continued refusal to add custodians that would address substantial deficiencies in CIR’s production of highly relevant categories of documents—documents that CIR has already admitted would be uniquely in the possession of OpenAI’s proposed custodians.

As background, OpenAI filed a motion on April 1 seeking an order compelling CIR to supplement its three individual custodians<sup>1</sup> with five additional custodians. ECF 250, 251. Each proposed custodian had been identified by CIR’s custodial Rule 30(b)(6) designee, Monika Bauerlein (CIR’s CEO), as possessing relevant documents not in the possession of CIR’s current custodians. *See* ECF 250 at 2–3. In opposing OpenAI’s motion, CIR repeatedly insisted that OpenAI first review CIR’s production for any deficiencies. Having now done so, OpenAI identified substantial deficiencies in the production of precisely the types of documents Bauerlein testified would be uniquely in the possession of OpenAI’s proposed custodians. While CIR has also insisted that it will manually collect missing documents, that cannot serve as a substitute for the identification of custodians and search terms ordered by the ESI Order, particularly in the face of broad production deficiencies. *See Thomas v. City of New York*, 336 F.R.D. 1, 5 (E.D.N.Y. 2020) (“self-collection [] is strongly disfavored,” particularly absent “steps to determine whether significant gaps exist”) (citation omitted).

During the May 27 conference, the Court ordered further conferral regarding CIR’s production deficiencies. Ex. A at 168:2-9. OpenAI subsequently sent CIR a detailed description of those deficiencies, along with explanations of why the missing documents would be in the possession of OpenAI’s proposed CIR custodians. Ex. B at 2–4. The parties remained at an impasse following their meet and confer by videoconference on June 5, and OpenAI renews its motion for the addition of the custodians below:

**Robert Wise (Director of Online Technology):** CIR agreed to produce *documents and communications* relating to its use of exclusion protocols on its websites (RFP 34); documents

---

<sup>1</sup> Those custodians were Monika Bauerlein, CEO; Madeleine Buckingham, CFO; and Khary Brown, VP of Media Sales and Business Development.

relating to the addition of copyright management information (“CMI”) to articles on its websites, such as documents about website build (RFP 101); historical versions of its robots.txt files (RFP 107); and documents and communications concerning modifications to its robots.txt files (RFP 108). ECF 251-6 at 8; ECF 251-7 at 2; ECF 251-1 at 4. These documents, none of which CIR claims are irrelevant, address, *inter alia*, whether CIR has conveyed an implied license by allowing the scraping of its content and whether CIR has properly included CMI with its content. To date, however, the only responsive documents from CIR are its current robots.txt files and a few isolated documents relating to its website build. It is inconceivable that CIR has no additional documents relating to its use of CMI, for example, when at least one of its publications, Reveal, changed its CMI practices in 2018. *See* Ex. C at 6–7. Further, Bauerlein testified that Wise is [REDACTED] and confirmed that [REDACTED] ECF 250-5 at 20:15-17, 184:6-185:4. As for CMI, Bauerlein again identified Wise as the individual [REDACTED] *Id.* at 169:12-171:16. Wise therefore likely possesses relevant documents not in the possession of current custodians.

**Marla Jones-Newman (VP for People and Culture):** CIR agreed to produce documents responsive to RFP 12 showing “[CIR]’s ownership of the asserted works,” including “employment and work-for-hire agreements”—the prerequisite for CIR’s copyright infringement claim. ECF 250 at 3. CIR has asserted at least 70,000-80,000 works, Am. Compl. Exs. 2–4, associated with at least 500 distinct authors based on OpenAI’s analysis of CIR’s production. However, CIR has produced less than 200 employment or freelance agreements, many of which are for individuals not associated with the asserted works. And while CIR’s asserted works go as far back as 1978, Am. Compl. ¶ 36, the produced agreements only date back to 2003. Thus, while CIR insists that it is manually collecting these agreements, that effort has clearly been insufficient in light of its substantial number of authors. Jones-Newman, who Bauerlein confirmed has the [REDACTED], likely possesses relevant documents not in the possession of current custodians. ECF 250-5 at 179:3-8.

**Emily Harris (Director of Finance):** CIR agreed to produce documents and correspondence relating to CIR’s analyses of “trends in readership or online subscriptions” (RFP 31), “the financial value of content” (RFP 32), and “the performance of content” (RFP 33), all of which go to the alleged harm, damages, and factor four of the fair use defense. ECF 250 at 3. While CIR has produced high-level analyses relating to advertising revenue and forecasts, its production is largely missing data underlying those analyses, analyses of the financial value of the content, or analyses of other types of revenue streams, such as licensing revenue. Moreover, less than 20% of the documents CIR has produced relating to financial analyses are for Reveal, including almost no substantive documents describing, for example, how Reveal arrived at projected numbers or the specific factors impacting Reveal’s revenue. While CIR insists that Bauerlein and Buckingham receive the final analyses and [REDACTED], Bauerlein testified that it was Harris [REDACTED] ECF 250-5 at 160:12-161:22, 162:1-13. Harris therefore likely possesses relevant documents not in the possession of current custodians.



**Ruth Murai (Research Director):** CIR agreed to produce documents responsive to RFPs relating to the originality of CIR's works (RFP 7), the use of third-party texts and sources in its works (RFPs 35–36), and plagiarism (RFP 57). ECF 251-6 at 2, 8; ECF 251-1 at 18. CIR also agreed to produce “documents regarding [CIR's] use of the Defendants' Gen AI tools in reporting [] of content.” ECF 250 at 3. These RFPs seek documents beyond formal internal or external complaints of plagiarism, which CIR has suggested Bauerlein would receive; they also seek day-to-day *determinations* and *conversations* around what constitutes acceptable use of third-party text and AI, which are highly relevant to ownership, CIR's own reliance on fair use, and OpenAI's unclean hands defense. Thus far, CIR has produced no responsive documents apart from its employee handbooks. Given that the majority of CIR's asserted works incorporate third-party material,<sup>2</sup> it is highly unlikely that CIR possesses no communications directed to the use of third-party text during the editing process for specific articles. And while CIR suggests that any plagiarism allegations would be raised with Bauerlein, CIR has produced no communications relating to plagiarism concerns for the 70,000-80,000 articles it has asserted. The absence of *any* such communications indicates that CIR's current collection methods are insufficient. Bauerlein also made clear that it is Murai, not Bauerlein, who [REDACTED] ECF 250-5 at 144:16-145:7. Murai therefore likely possesses relevant documents not in the possession of current custodians.

**Michael Mechanic (Senior Editor):** CIR agreed to produce documents relating to the use of third-party texts and sources (RFPs 35–36), and documents *and communications* relating to requests and licenses to third parties (RFPs 51–52) and content-sharing licensing programs (RFP 72). ECF 251-6 at 8, 15; ECF 251-1 at 18. CIR does not dispute that all such documents are relevant to fair use, OpenAI's unclean hands defense, and the value of CIR's works. Despite CIR's routine use of third-party material in its asserted works, it has produced documents describing only a handful of instances where CIR made a request for such use. CIR has also produced no documents relating to content-sharing programs with third parties, aside from a few isolated emails and portions of agreements relating to two (out of the fourteen) partners of its Climate Desk syndication project. While CIR claims that it is manually collecting “publication agreements,” Ex. A at 166:8-10, manual collection cannot conceivably cover all responsive *documents and communications*, which (as CIR does not dispute) are relevant to the licensing value of CIR's works. Further, Bauerlein testified that Mechanic, rather than herself or other current custodians, would [REDACTED] ECF 250-5 at 137:22-142:22. Mechanic therefore likely possesses relevant documents not in the possession of current custodians.

For the reasons above, the Court should compel CIR to add the aforementioned custodians.

<sup>2</sup> See, e.g., Kevin Drum, *Is America Just an Oversized Ireland?* (May 4, 2010), <https://www.motherjones.com/kevin-drum/2010/05/america-just-oversized-ireland/> (containing a book review excerpt).



Respectfully,

KEKER, VAN NEST &  
PETERS LLP\*

LATHAM & WATKINS  
LLP

MORRISON &  
FOERSTER LLP\*

/s/ Andrew S. Bruns  
Andrew S. Bruns

/s/ Herman H. Yue  
Herman H. Yue

/s/ Rose S. Lee  
Rose S. Lee

cc: All Counsel of Record Line (via ECF)

---

\* All parties whose electronic signatures are included herein have consented to the filing of this document.